

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

**IN RE DIDI GLOBAL INC. SECURITIES
LITIGATION**

This Document Relates To: All Actions

Master Docket

Case No. 1:21-cv-05807-LAK-VF

**STIPULATED PROTOCOL
FOR DISCOVERY OF
ELECTRONICALLY STORED
INFORMATION (“ESI
PROTOCOL”)**

CLASS ACTION

Lead Plaintiff Junhong Cao and named plaintiffs Shereen El-Nahas, Alaka Holdings Ltd., Bosco Wang, Daniil Alimov, and Njal Larson (together, “Plaintiffs”), Defendant Didi Global Inc. (“Didi” or the “Company”), Defendants Goldman Sachs (Asia) L.L.C. (“Goldman Sachs”), Morgan Stanley & Co. LLC (“Morgan Stanley”), J.P. Morgan Securities LLC (“J.P. Morgan”), BofA Securities, Inc. (“BofA Securities”), Barclays Capital Inc. (“Barclays”), Citigroup Global Markets Inc. (“Citigroup”), HSBC Securities (USA) Inc. (“HSBC”), UBS Securities LLC (“UBS”), Mizuho Securities USA LLC (“Mizuho”), and China Renaissance Securities (US) Inc. (“CRS”) (together, the “Underwriter Defendants), and Defendants Will Wei Cheng, Jean Qing Liu, Stephen Jingshi Zhu, Alan Yue Zhuo, Adrian Perica, Kentaro Matsui, Martin Chi Ping Lau, Zhiyi Chen, and Daniel Yong Zhang (together, the “Individual Defendants”) (collectively, “Defendants,” and together with Plaintiffs, the “Parties” and each a “Party”), by and through their respective undersigned counsel, hereby stipulate and agree to the following Stipulated Protocol for Discovery of Electronically Stored Information (the “ESI Protocol” or “Protocol”).

I. GENERAL PROVISIONS

1. This Protocol will govern discovery of ESI and hard copy documents in this case as a supplement to the Federal Rules of Civil Procedure, the Local Rules of the United States District Courts for the Southern and Eastern Districts of New York (the “Local Rules”), and the individual Policies and Procedures of District Judge Lewis A. Kaplan¹ and Magistrate Judge Valerie Figueredo² (together, the “Individual Rules”), and any other applicable orders and rules.

2. The Parties are aware of the importance the Court places on cooperation and commit to cooperate in good faith throughout this matter. The Parties shall follow Local Rule 26.4 and the Courts’ Individual Rules with respect to any and all discovery disputes.

3. Unless otherwise agreed to or directed by the Court, each Party will bear the costs of collecting, processing and reviewing documents in that Party’s possession, custody, or control according to this Protocol.

4. Nothing in this Protocol shall be construed to affect the discoverability or admissibility of any document. All objections to the discoverability or admissibility of any document are preserved and may be asserted at any time in accordance with the applicable Federal Rules of Evidence, Federal Rules of Civil Procedure, and/or Local Rules or Individual Rules of this Court.

5. Each production of ESI shall be accompanied by correspondence indicating the Bates range and any applicable passwords.

¹ Available at:

https://www.nysd.uscourts.gov/sites/default/files/practice_documents/LAK%20Kaplan%20Individual%20practices%20_%204.18.2020.pdf

² Available at:

https://www.nysd.uscourts.gov/sites/default/files/practice_documents/VF%20Figueredo%20Rules%20of%20Practice%20in%20Civil%20Cases%20as%20of%209.19.23.pdf

II. DEFINITIONS

1. The Parties expressly adopt and incorporate by reference (as if fully set forth herein) the uniform definitions set forth in Local Rule 26.3. Additional defined terms are set forth below. In the event of any conflict, the more specific definitions set forth below shall control.

2. Action: means the above-referenced action.

3. Document: has the meaning ascribed to it in Federal Rule of Civil Procedure 34(a)(1)(A) and S.D.N.Y. Local Rule 26.3. The term “Document” shall include Hard-Copy Documents and ESI as defined herein.

4. Email: means electronic messages sent or received over the internet using simple mail transfer protocol or internally within an organization’s mail service, including, but not limited to, Microsoft Exchange.

5. Email Threading: means a single email conversation that starts with an original email (the beginning of the conversation) and includes all subsequent replies and forwards pertaining to that original email.

6. ESI: an abbreviation of “electronically stored information,” which has the meaning ascribed to in Federal Rule of Civil Procedure 34(a)(1)(A).

7. Extracted Text: means text extracted from a Native Format file and includes at least all headers, footers, document body information, and any hidden text, if available. The extracted text shall contain the content of any hypertext markup language (“HTML”) present in the Native Format file, specifically including (but not limited to) the full anchor and target text of any Hyperlinks, as defined herein. The extracted text must not include text of characters that were not part of the text of the original Native Format file, including but limited to, Bates Numbers and Endorsements (except in the cases of redactions).

8. Go-Get: means a request to manually retrieve and produce a specifically-identified document, or small number of documents.

9. Hard-Copy Document: means Documents existing in paper form at the time of collection.

10. Hash Value: means the unique numerical identifier that can be assigned to a file, a group of files, or a portion of a file, based on a standard mathematical algorithm applied to the characteristics of the data set. The most commonly used algorithms, known as MD5 and SHA, will generate numerical values so distinctive that the chance that any two data sets will have the same Hash Value, no matter how similar they appear, is less than one in one billion. “Hashing” is used to guarantee the authenticity of an original data set and can be used as a digital equivalent of the Bates stamp used in Hard-Copy Document productions.

11. Hyperlink: means an HTML link from an electronic Document to another electronic Document, typically activated by clicking on a highlighted word or image on the screen, and includes both the “anchor” text (*i.e.*, the ordinarily visible, clickable text of the hyperlink, often presented in blue in the linking document) and the “target” text (*i.e.*, the ordinarily invisible hypertext containing the Uniform Resource Locator (“URL”) for the location of the linked document).

12. Instant Messages: means real-time communications sent via chat client, short messaging service (“SMS”), Multimedia Messaging Service (“MMS”), or Rich Messaging Service (“RMS”), including but not limited to: Slack, Microsoft Teams, Google Talk, Google Chat, Google Hangouts, Skype, Facebook Messenger, Instagram Messenger, Line, Wire, Signal, Telegram, Snapchat, Wickr, Discord, WhatsApp, Bloomberg Messenger, WeChat, DingTalk, Tencent QQ, Sina Weibo, TikTok (Douyin), or any similar or proprietary instant messaging system.

13. Load File: means an electronic file that is used to import all required production information into a document database, including, if available, document images, Extracted Text or OCR text, Native Format files where required by this ESI Protocol, and Metadata, as well as information indicating document breaks, and document relationships such as those between an Email or Instant Message and its attachments, and a document and information related to embedded content.

14. Metadata: means (i) structured information about ESI that is created by the file system or application that is not ordinarily viewable or printable from the application that generated, edited, or modified such native file, embedded in the Document or Email and sometimes modified through ordinary business use; (ii) information generated automatically by the operation of a computer or other information technology system when a native file is created, modified, transmitted, deleted, or otherwise manipulated by a user of such system; (iii) information, such as Bates Numbers, created during the course of processing Documents or ESI for production; and (iv) information collected during the course of collecting Documents or ESI, such as the name of the Media device on which it was stored, or the custodian or non-custodial data source from which it was collected. Metadata of the ESI describes, *inter alia*, the characteristics, origins, usage, and validity of the collected ESI.

15. Native Format: means the format of ESI in the application in which such ESI was originally created.

16. OCR: means the optical character recognition technology used to read Hard-Copy Documents or electronic images of Documents and output such Documents to a searchable text format.

17. Producing Party: means any Party or Third Party in the Action that produces Documents.

18. Protective Order: means the Confidentiality Stipulation and Protective Order, Dkt. No. 190, entered in this Action on May 22, 2029, as amended by any subsequent order(s) of the Court.

19. Receiving Party: means a Party in the Action to whom Documents are produced.

20. Responsive Document: means any Document that is responsive to any discovery request or subpoena served on the Producing Party in the Action and which the Producing Party has agreed or been ordered to produce, subject to the limitations set forth in the Federal Rule of Civil Procedure and/or Court order.

21. Tagged Image File Format or TIFF: refers to the CCITT Group IV graphic file format for storing bit-mapped images of ESI or Hard-Copy Documents.

22. Third Party: Any other person/entity, other than a party to this Action.

III. PRESERVATION AND IDENTIFICATION OF POTENTIALLY RELEVANT INFORMATION AND ESI

The Parties acknowledge that they have an obligation to take reasonable and proportional steps to preserve discoverable information in the Party's possession, custody, or control. Nothing in this ESI Protocol shall affect the Parties' respective preservation obligations imposed by rule or law.

IV. SEARCH, COLLECTION & REVIEW

The Parties will meet and confer in good faith to discuss any methodologies to be employed for the collection and production of documents in response to any Fed. R. Civ. P. 34 request for production, such as the use of search terms, date ranges, electronic sources, or other filters. The

Parties may use advanced search, analytical, and retrieval technologies like predictive coding or other technology-assisted review (“TAR”) in order to prioritize the review of or the grouping of documents requested or returned from the use of search terms, date ranges, electronic sources, or other filters as agreed to between the Parties. Advanced search, analytical, and retrieval technologies such as predictive coding or other TAR shall not be used to exclude from production any documents returned from search protocols agreed to between the Parties. The Producing Party, to the extent search terms are used, will upon request undertake reasonable efforts to provide search term reports identifying the “hit” count of documents and documents including family per search term, the unique hit count of documents and documents including family per search term, and the total number of unique documents that hit on the search term list as a whole.

V. PRODUCTION FORMATS FOR DOCUMENTS FROM HARD COPY DOCUMENTS

A. Generally: Except for redacted and logged documents as described in Section 0 of this ESI Protocol, *infra*, the Parties will produce documents originating from Hard Copy Documents in CCITT Group IV single-page TIFF format (black and white, 300 dpi) If an original document contains color, it need not be produced in color in the first instance. However, the Producing Party will honor reasonable requests for a color image, if color is necessary to understand the meaning or content of the document. Color documents shall be produced as single-page, 300 DPI JPG images with JPG compression and a high-quality setting to as to preserve legibility. The Parties agree not to degrade the searchability or legibility of documents as part of the document production process.

B. Document Unitization: In scanning paper Hard Copy Documents, each page of paper should be output to a single page TIFF file. Distinct, logical document breaks should be

defined as such in a standard load file as described herein. In the case of an organized compilation of separate Documents (*e.g.*, a binder containing several separate Documents behind numbered tabs) the Document behind each tab should be scanned separately, but any document or family relationship among the scanned Documents in the compilation should be reflected in the data load file at the appropriate standard fields. Pages containing post-it notes or other detachable notes that obscure the underlying Document should be scanned once with the detachable note intact, and then again without it, and made part of the same Document. The Parties shall make reasonable efforts to unitize the Documents in accordance with this paragraph, but recognize that certain documents and categories of documents might not reasonably be scanned to comply.

C. Load File: The Parties will provide a standardized load file compatible with Relativity and with a Bates number field included in the load file to match text and metadata with TIFF/JPJ images. With respect to hard copy documents, data on the load file will include the following metadata or their equivalents (when available), as defined in Section VI of this ESI Protocol:

- a. BEGDOC;
- b. ENDDOC;
- c. BEGATTACH;
- d. ENDATTACH;
- e. CUSTODIAN;
- f. CONFIDENTIALITY;
- g. PGCOUNT;
- h. REDACTED; and
- i. TEXTLINK.

VI. PRODUCTION FORMATS FOR ESI

A. Production Format Generally.: Except for redacted and logged documents as described in Section 0 of this ESI Protocol, *infra*, the Parties will produce documents originating as ESI in TIFF format, with the exception of spreadsheets, audio, and video files, which shall be produced in native format. ESI shall be produced with extracted text, along with the below listed metadata fields, where reasonably available. If an original document contains color, it need not be produced in color in the first instance. However, the Producing Party will honor reasonable requests for a color image, if color is necessary to understand the meaning or content of the document. Color documents should be produced as single-page, 300 DPI JPG images with JPG compression and a high-quality setting as to not degrade the original image. TIFFs/JPGs will show any and all text and images, including track changes where applicable, that would be visible to the reader using the native software that created the document. For example, TIFFs/JPGs of email messages should include the BCC line. If the image does not accurately reflect the document as it was kept in the usual course of business, including all comments, edits, tracking, etc., the Parties agree to meet and confer in good faith on production format options. Subject to the caveat below, embedded files shall be produced as attachments to the document that contained the embedded file, with the parent/child relationship preserved. The embedded files will be marked with a "YES" in the load file under the "EMBEDDED" metadata field. The Parties agree that embedded documents need not be extracted as separate documents as long as they are displayed in their entirety and legible in the parent document (e.g., logos, icons, or images). Compressed file types (e.g., .ZIP, .RAR, .CAB, .Z) should be decompressed so that the lowest level document or file is extracted. The Parties will provide a standardized load file compatible with Relativity.

B. Metadata: Each of the Metadata and coding fields set forth below that can reasonably and technically be extracted or generated from a Document shall be produced for that Document. All ESI will be collected in a manner that preserves the associated metadata intact (specifically including, but not limited to, the below-listed fields for “Created Date/Time” and “Last Modified Date/Time”). To the extent that metadata below does not exist or is not reasonably accessible or available for any documents produced, nothing in this Order shall require any Party to extract, capture, collect, or produce such data, with the exception of the following fields, if available: (a) BegBates; (b) EndBates; (c) BegAttach; (d) EndAttach; (e) All Custodian; (f) Redacted (Y/N); (g) Confidentiality; and (h) HashValue. Metadata shall be provided in a Concordance-format delimited file with a .DAT file extension and ASCII 020 and 254 delimiters for column break and text qualifier. The first line shall be the header with field names, and each subsequent line shall contain the fielded data for each Document. The Parties also agree that every production of previously unproduced Documents, the Metadata field values for each Metadata field header will be mapped to the same Metadata field source, unless otherwise agreed in writing. For redacted items which were originally ESI, all non-privileged metadata fields will be provided and will include all non-privileged, non-redacted data. Redacted documents shall be identified as such in the load file provided with the production. Nothing herein shall require a Producing Party to create or produce Metadata that does not exist, cannot be generated by ESI processing software, or is not reasonably or technically collectable. The Parties agree, in particular, that “go-get” collections of small sets of documents do not require forensic-style collection to preserve all metadata. When the Producing Party is unable to produce metadata for a particular field or ESI document, the Receiving Party may request an explanation of that inability. The Parties shall then

meet and confer to attempt to resolve the problems. With respect to ESI, data on the load file will include the following metadata fields or their equivalents (when reasonably collectable):

Field Name	Description
BEGDOC	Begin Document Bates number
ENDDOC	End Document Bates number
BEGATTACH	Begin Bates number of family unit
ENDATTACH	End Bates number of family unit
CUSTODIAN	Individual from whom the document originated
ALL CUSTODIAN	Name of the agreed-upon custodians, in addition to the primary custodian, from whose files the item would have been produced if it had not been de-duplicated
FROM	Author of the email message
TO	Recipient(s) of the email message
CC	Recipient(s) of "Carbon Copies" of the e- mail message
BCC	Recipient(s) of "Blind Carbon Copies" of the email message
DOCUMENT SUBJECT	Subject field extracted from the metadata of the native file
EMAIL SUBJECT	Subject field extracted from the metadata of an email file
DOCUMENT TITLE	The title of a document
DOCUMENT AUTHOR	Author field extracted from the metadata of the native file
FILE EXTENSION	File extension of a document
CREATED DATE/TIME	Date and time file was created
LAST MODIFIED DATE/TIME	Date and time document was last modified
EMAIL SENT DATE/TIME	Date and time email was sent

EMAIL RECEIVED DATE/TIME	Date and time email was received
DATE/TIME APPOINTMENT START	Start date and time of calendar appointment entry
DATE/TIME APPOINTMENT END	End date of calendar appointment entry
LAST EDITED BY	Name of the last person to edit the document from extracted metadata
ACCESSED BY	Name of each person who accessed the document and the times(s) each such person first and last accessed the document
PGCOUNT	Number of pages in a document (image records)
FILE SIZE	File size in bytes
LOGICALPATH	Original logical path location of the document in the repository from which it was collected.
FILE TYPE	File Type: email, image, spreadsheet, presentation, etc.
MD5 HASH	Unique identifier of the file
FILE NAME	Original name of the file or subject of email, including file extension
CONFIDENTIALITY	User-generated field that will indicate confidentiality pursuant to the Protective Order (i.e., CONFIDENTIAL). Otherwise, blank.
REDACTED	YES if applicable. Otherwise, blank.
TECHNICAL ISSUE	YES where slip sheet reads "Technical issue—file cannot be processed." Otherwise, blank.
TEXTLINK	Contains Path to .TXT files
NATIVELINK	Contains Path to native files
SOURCE	Physical location where documents were collected (e.g., Custodian's laptop, hard drive, shared drive, cell phone)

ISEMBEDDED	Identifies if a document is embedded in another document.
------------	---

C. Reproduction in Native Format: The Parties further agree that to the extent any Party seeks production in native format of specifically identified ESI produced originally in TIFF form for good cause shown, the Producing Party shall respond reasonably and in good faith to any such request.

VII. MISCELLANEOUS PROVISIONS

A. Additional Custodians or Data

The Parties will endeavor to seek documents or data from all relevant repositories of information during the parties' initial discussions on the parameters of such discovery, but reserve the right to request inclusion of additional custodians or non-custodial data sources whose relevance was discovered via documents or data produced, or testimony given, or for other good cause shown, within a reasonable time of such discovery, subject to the Local Rules and Individual Rules, through the close of fact discovery. If the Producing Party objects to the inclusion of such sources, the Parties will meet and confer to resolve the matter. If the Parties cannot reach resolution, the Court will determine the matter.

B. Additional Formatting Processes

Any Producing Party may De-NIST their electronic collections before searching and review using the then-current NIST list published by the National Institute of Standards and Technology.

To the extent that relevant ESI cannot be rendered or reviewed without the use of proprietary or special software, the Parties shall meet and confer to minimize any expense or

burden associated with the production of such documents or information in an acceptable format, including issues as may arise with respect to obtaining access to any such software.

The Parties will make reasonable efforts to ensure that all encrypted or password-protected ESI are successfully processed prior to production. To the extent encrypted or password-protected ESI are successfully processed, the Parties have no duty to identify further the prior encrypted status of such documents. To the extent security protection for such ESI cannot be successfully processed despite reasonable efforts, the Parties shall not be required to produce them. If a member of a document family that has otherwise been determined to be responsive cannot be technically processed (e.g., unsupported file format, file corruption, inaccessible password-protected document) using reasonable efforts, those technical problems shall be identified and disclosed to the receiving party by production of a Bates numbered slip sheet that states “Technical issue—file cannot be processed.” A receiving Party thereafter may raise with the Producing Party any questions or concerns, and the Parties shall meet and confer to attempt to resolve any issues.

C. Bates Numbering

The Parties shall assign a Bates number to individual pages of TIFF documents and a Bates number to each document produced in native format. Bates numbers shall be unique across the entire document production and sequential within a given document. All volumes of production shall be identified by the same Bates prefix and a numerical sequence. Re-produced documents shall include their original Bates numbers, if any.

D. Confidentiality Designations

If a document or other ESI has a confidentiality designation, the designation shall be stamped on the face of all TIFF images pertaining to such document pursuant to the Protective Order. If the receiving Party believes that a confidentiality designation obscures the content of a

document, then the receiving Party may request that the document be produced with the confidentiality designation in a different position. No Party may attach to any filing or any correspondence addressed to the Court (including any Magistrate Judge), or any adverse or third party, or submit as an exhibit at a deposition or any other judicial proceeding, a copy (whether electronic or otherwise) of any native format document produced by any party without ensuring that the corresponding Bates number and confidentiality legend, as designated by the Producing Party, appears on the Document or in the filename, if used in native format. For each document that is marked confidential, a confidentiality field will be populated with the word “Confidential” in the .DAT file.

E. De-Duplication and Document Families

The Parties will perform de-duplication of ESI within and across custodians according to MD5 or SHA-1 hash values at a family level and will produce only a single copy of identical ESI, except where de-duplication would break up document families. De-duplication shall not break up document families. The custodian associated with the first copy of a document processed will be considered the primary custodian for that document (the custodian who will be used as the basis for determining which other collected documents are duplicates). The Producing Party shall produce an “All Custodians” metadata field with each production as specified in Section VI above which lists every custodian or source who or which possessed a duplicate document. If the Parties de-duplicate ESI, they shall provide custodian associations in a semi-colon delimited field that includes duplicate custodian name information. If additional documents are produced after substantial completion of document discovery, the Producing Party shall produce an overlay that provides an updated “All Custodian” field at the time new productions are made.

If there are any handwritten notes, or any other markings, on a document, it shall not be considered a duplicate. Any document that contains an alteration, marking on, or addition to the original document shall be treated as a distinct version, and shall be produced as such. These alterations include, but are not limited to, handwritten notes, electronic notes/tabs, edits, highlighting or redlining. If such markings/alterations are made in color, the documents must, to the extent reasonably practicable, be produced in color as originally intended.

For purposes of production, and to the extent feasible on an automated, scalable basis with existing technology, responsive documents should include all parent items (mail, calendar, contacts, tasks, notes, etc.) and child files (attachments of files to email to internal or nonpublic documents, or other items), with the parent/ child relationship preserved.

F. Hyperlinks

Plaintiffs' Position

- Plaintiffs' Proposed Language: For Documents responsive to properly propounded requests issued pursuant to Fed. R. Civ. P. 34 or 45, the Producing Party shall conduct an automated search across all emails to be produced, where the Producing Party has access to a tool capable of performing such an automated search, to identify any documents that contain hyperlinks to an internal or non-public document. Additionally, the Receiving Party may make reasonable and proportionate requests that the Producing Party conduct a reasonable search for internal or non-public documents identified via hyperlinks. For each hyperlink identified by the Producing Party or the Receiving Party, the Producing Party shall conduct a reasonable search for the document corresponding with the hyperlink. The Producing Party shall process

Case No. 1:21-cv-05807-LAK-VF

STIPULATED PROTOCOL FOR DISCOVERY OF ELECTRONICALLY STORED INFORMATION ("ESI PROTOCOL")

and produce relevant Documents corresponding with the hyperlinks with reasonably available metadata, and each such hyperlinked Document shall be treated as a traditional attachment to the original Document containing the hyperlink and be produced in the same manner as other document families (i.e. with corresponding metadata linking family members in the load file). To the extent possible, the hyperlinked document produced should be the contemporaneous document version that was likely present at the time the underlying Document was sent. If a Producing Party identifies any technical obstacle to the collection or production of hyperlinked files included in (or referenced to) any specific repository, the Producing Party will provide the Requesting Party a description of the nature of the issue, and the Parties will meet and confer in good faith to negotiate a mutually-agreeable compromise.

- **Plaintiffs' Explanation:** Modern corporate communications often contain hyperlinks to documents stored on corporate servers, in the cloud, or on other shared repositories. Hyperlinks have, in many cases, replaced the practice of attaching Word documents or PDFs to emails. These hyperlinked documents are essentially attachments, and Plaintiffs' proposed language would treat them as such: if a responsive, non-privileged document contains a hyperlink, the hyperlinked document should be produced and linked to the original document via metadata to preserve the parent/child relationship. Plaintiffs' proposed language would require Producing Parties – only to the extent such an automated search is feasible – to run an automated search to identify responsive documents that contain hyperlinks, locate the hyperlinked documents, and

produce them. Additionally, to the extent any hyperlinked document is missed by a Producing Party, the Receiving Party would be able to send specific requests for hyperlinked documents to be produced. Finally, the parties would meet and confer concerning any technological obstacles to such automated searches. Plaintiffs' proposal is modeled on clauses from other ESI protocols so-ordered in this district. See *In re Google Digital Advert. Antitrust Litig.*, 2023 WL 2557433, at *9 (S.D.N.Y. Mar. 17, 2023); *In re Aceaminophen - ASD-ADHD Prod. Liab. Litig.*, No. 22MD3043 (DLC), 2023 WL 196157, at *5 (S.D.N.Y. Jan. 17, 2023); see also *In re Uber Techs., Inc., Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *7 (N.D. Cal. Apr. 23, 2024) ("Attachment(s)' shall be interpreted broadly and includes, e.g., traditional email attachments and documents embedded in other documents (e.g., Excel files embedded in PowerPoint files) as well as modern attachments, pointers, internal or non-public documents linked, hyperlinked, stubbed or otherwise pointed to within or as part of other ESI (including but not limited to email, messages, comments or posts, or other documents)."); Lea Malani Bays and Stuart A. Davidson, *The Missing Links: Why Hyperlinks Must Be Treated as Attachments in Electronic Discovery*, 92 U. Cin. L. Rev. 979 (2024).³ An automated search does not impose undue burden on Defendants. And it is premature for Defendants to claim such an automated search is not technologically possible on their specific systems given that Defendants have

³ Available at: <https://scholarship.law.uc.edu/uclr/vol92/iss4/9>
Case No. 1:21-cv-05807-LAK-VF

refused to even identify their systems to Plaintiffs, much less attempted to automate the task. *C.f. In re Stubhub Refund Litig.*, 2024 WL 2305604, at *1 (N.D. Cal. May 20, 2024) (finding that automating production of hyperlinks was not feasible only after the defendant “spent hundreds of hours trying to find linked documents and despite having retained an outside e-discovery vendor to assist with this effort”).

• **Defendants’ Position**

- ~~Defendants’ Proposed Language~~ After receiving a production of documents responsive to properly propounded requests issued pursuant to Fed. R. Civ. P. 34 or 45, the Receiving Party may make reasonable and proportionate requests that the Producing Party conduct a reasonable search for internal or non-public document identified via hyperlinks.

~~**Defendants’ Explanation.** Plaintiffs’ proposal—which would require Defendants to hunt down the source of every hyperlink in its ESI production—impermissibly expands Defendants’ obligations in a way that is neither technologically possible nor reasonable and proportionate to the needs of the case. To justify imposing a substantial burden on Defendants, Plaintiffs insist that “hyperlinked documents are essentially attachments.” But that is not true. *See In re Stubhub Refund Litig.*, 2024 WL 2305604, at *2 (N.D. Cal. May 20, 2024) (“linked-to documents are *not* attachments”); *Nichols v. Noom Inc.*, 2021 WL 948646, at *4 (S.D.N.Y. Mar. 11, 2021) (“[T]he Court does not agree that a hyperlinked document is an attachment.”). Email systems have long permitted users to link to documents. Defendants do not dispute that hyperlinks~~

~~provide indicia of potentially relevant and discoverable material, but the more appropriate way to gather those materials is to search and produce directly from their sources. As part of their discovery obligations, Defendants already search for locations likely to have responsive materials and collect from them, including cloud-based repositories. If Plaintiffs wish to have hyperlinked documents produced from such repositories matched to an original document containing a link, Plaintiffs can do so based on each document's metadata. But, to the extent that Plaintiffs insist on creating an artificial parent/child relationship, Defendants and courts agree that it is not proper. *See Nichols*, 2021 WL 948646, at *4. Nor is it technologically possible the majority of times based on current commercially available systems. *See In re Stubhub*, 2024 WL 2305604, at *1-2 (evaluating expert declarations and concluding that ESI protocol's "hyperlink requirement is technologically impossible to fulfill most of the time"). Consistent with the Federal Rules, Defendants will search and produce emails and custodial files as part of discovery negotiations. Further, as reflected in Defendants' proposal, they will entertain reasonable and proportionate requests to search for hyperlinked documents identified by Plaintiffs.~~

G. Encryption. The Parties will make reasonable efforts to ensure that all encrypted or password-protected Documents are successfully processed for review and production under the requirements of this Order. Upon request from either Party, the Parties shall meet and confer in good faith regarding reasonable efforts or mechanisms to remove such security protection or the production of available Metadata.

H. Responsive ESI Discovered Outside of Search Methodology. Notwithstanding any of the foregoing, if any Party identifies ESI or Documents that are responsive and subject to discovery in this matter outside of the search methodology(ies) agreed to by the Parties, it shall produce all such Documents except any that must be withheld or redacted on the basis of privilege, work product, or other grounds..

I. ESI Not Required. The Parties agree that the following ESI does not require preservation, collection, review, or production because they are either not reasonably accessible or because the remote possibility of additional relevant information existing in not reasonably accessible sources is substantially outweighed by the burden and cost of preservation, collection, review, and production, provided, however, that nothing herein shall prevent a Party from subsequently requesting that Documents from the sources identified below be preserved and produced if specific facts demonstrate a particular need for such Documents that justifies the burden of preservation and retrieval:

- a. Deleted, shadowed, damaged, residual, slack, fragmented, or other data on hard drives, solid-state drives (SSDs), or similar physical media, and which are only accessible by forensics;
- b. Random access memory (RAM), temporary files, or other ephemeral data that is difficult to preserve without disabling the operating system;
- c. On-line access data such as temporary Internet files, history, cache, cookies, and the like;
- d. Data stored on photocopiers, scanners, and fax machines;
- e. Data in metadata fields that are frequently updated automatically, such as last-opened dates;

f. Data maintained or duplicated in any electronic backup system for the purpose of system recovery or information restoration, including, but not limited to, system recovery backup tapes or other media, continuity of operations systems, and data or system mirrors or shadows, if such data are routinely purged, overwritten, or otherwise made not reasonably accessible in accordance with an established routine system maintenance policy;

g. Server, system, or network logs;

h. Data remaining from systems no longer in use that is unintelligible on the systems in use;

i. Video surveillance data; and

j. Email, calendars, and contact data sent to or from mobile devices (*e.g.*, iPhone, iPad, Android, and Blackberry devices), provided that a copy of all such electronic data is saved elsewhere (such as on a server, laptop, desktop computer, or “cloud” storage).

J. Preservation of Parent-Child Relationships: The Parties agree that if any part of an email or its attachments is responsive, the entire family of email and attachments will be produced, to the extent feasible on an automated, scalable basis with existing technology, except any family members that must be withheld or redacted on the basis of privilege, work product, or other grounds. The Parties shall use reasonable efforts to preserve parent-child relationships (the association between an attachment and its parent document). The Parties will provide a BEGATTACH and ENDATTACH for each produced attachment in the load file. If any member of a document “family” (*e.g.*, an email attachment) is withheld as privileged, then the producing party shall produce a one-page slipsheet TIFF image stating, “Withheld on Privilege Grounds”, and bearing a unique Bates number.

K. RedactionsPlaintiffs' Position

- Plaintiffs' Proposed Language: Except as set forth herein, responsive documents shall not be withheld or redacted on any basis other than the attorney-client privilege, attorney work product doctrine, or other recognized protection or privilege under U.S. law. In addition, a Party may redact the following personally sensitive or identifiable information: bank and brokerage account numbers; social security numbers; personal addresses; personal telephone numbers; other personal identifying information; and the names of minor children. For redacted items that were originally ESI, all metadata fields noted in this Protocol that do not contain privileged or protected information will be provided and will include all non-redacted data. The Parties specifically agree that documents produced shall not be redacted for relevance, responsiveness, or any other basis not contemplated herein. The Parties will produce redacted documents in TIFF format with corresponding searchable text and associated metadata for the documents, ensuring the redacted content is fully protected from disclosure. To the extent spreadsheets contain redacted text, the Parties shall provide a copy of the native file with redactions. For redacted items that were originally ESI, all metadata fields noted in this ESI Protocol that do not contain privileged or protected information will be provided and will include all non-redacted data. If, during the course of discovery, the Parties identify other kinds of information that a Party has

reasonable basis for redacting, the Parties will meet and confer regarding it before such redactions are made.

- **Plaintiffs' Explanation:** Consistent with the Court's directives at the May 22, 2024 conference, Plaintiffs believe that Defendants should not be allowed to redact or withhold documents based on PRC privacy or secrecy laws. See 5/22/24 Tr. at 5:6-11 ("I understand that, but you also need to understand, and I think I've had occasion to write on this previously, that is the U.S. rules that apply. If a foreign party invokes blocking statutes and the like, it will not stop me from ordering discovery and a timetable that suits the needs of this case in this court."). Additionally, Plaintiffs do not believe the parties should be permitted to redact information from relevant, responsive documents simply because the producing party unilaterally deems that information to be irrelevant or nonresponsive. See *Christine Asia Co. v. Alibaba Grp. Holding Ltd.*, 327 F.R.D. 52, 54 (S.D.N.Y. 2018) ("The weight of authority in this Circuit goes against allowing a party to redact information from admittedly responsive and relevant documents based on that party's unilateral determinations of relevancy."); *In re State St. Bank & Tr. Co. Fixed Income Funds Inv. Litig.*, 2009 WL 1026013, at *1 (S.D.N.Y. Apr. 8, 2009) ("such redactions are generally unwise" because they "breed suspicions, and they may deprive the reader of context").

Defendants' Position

Defendants' Proposed Language: A Party may redact the following personally sensitive or identifiable information: bank and brokerage account

Case No. 1:21-cv-05807-LAK-VF

STIPULATED PROTOCOL FOR DISCOVERY OF ELECTRONICALLY STORED INFORMATION ("ESI PROTOCOL")

numbers; social security numbers; personal addresses; personal telephone numbers; other personal identifying information; and the names of minor children. For redacted items that were originally ESI, all metadata fields noted in this Protocol that do not contain privileged or protected information will be provided and will include all non-redacted data. The Parties will produce redacted documents in TIFF format with corresponding searchable text and associated metadata for the documents, ensuring the redacted content is fully protected from disclosure. To the extent spreadsheets contain redacted text, the Parties shall provide a copy of the native file with redactions. For redacted items that were originally ESI, all metadata fields noted in this ESI Protocol that do not contain privileged or protected information will be provided and will include all non-redacted data. If, during the course of discovery, the Parties identify other kinds of information that a Party has a reasonable basis for redacting, the Parties will meet and confer regarding it before such redactions are made.

~~**Defendants' Explanation:** As an initial matter, Plaintiffs' proposed provision ignores that a party may withhold responsive documents on any of the bases identified in this ESI Protocol, the Federal Rules of Civil Procedure, the Federal Rules of Evidence, or other applicable federal or state law. Additionally, Defendants believe that any dispute about the issue of theoretical document transfer restrictions imposed by PRC law is premature, as Defendants have not withheld or redacted any documents on the basis of PRC laws, nor have they made any representation that they currently plan to do so. Regardless, even if~~

Case No. 1:21-cv-05807-LAK-VF

STIPULATED PROTOCOL FOR DISCOVERY OF ELECTRONICALLY STORED INFORMATION ("ESI PROTOCOL")

~~this issue does arise, Defendants believe that it is more appropriate to deal with it at that time and with reference to the actual document at issue, so that the court can undertake the required comity analysis in the context of a concrete dispute, rather than as a theoretical matter. See *Wultz v. Bank of China Ltd.*, 910 F. Supp. 2d 548, 552-53 (S.D.N.Y. 2012) (describing the factors that courts in the Second Circuit must consider when “evaluating the propriety of an order directing the production of information or contravention of foreign law”).~~

L. Privilege Log: The Parties agree that for all documents and information that are withheld or redacted on privilege, work product, or any other grounds⁴, the following information will be provided in an easily readable chart (“Privilege Log”):

- a. Privilege log reference number;
- b. Bates range of document (if applicable);
- c. Author(s) (designating which, if any, are attorneys);
- d. Recipient(s), including copyee(s) and blind carbon copyee(s) (designating which, if any, are attorneys);
- e. The file name, for non-email documents;
- f. The email subject, for email documents;
- g. The custodian;
- h. The sent date of email documents;
- i. The creation date, for non-email documents;
- j. The file type;

⁴ Nothing in this provision shall be interpreted as any Party waiving its rights to object to the propriety of a Producing Party withholding or redacting information on any grounds.

- k. Categorical description of the documents in a manner that, without revealing information itself privileged or protected, will enable other parties to assess the claim; and
- l. The asserted basis for withholding.

The following categories of documents need not be logged:

- a. Documents that reflect communications with the undersigned counsel or members of their law firms that are related to this litigation, and that post-date the filing of the initial complaint in this Action; and
- b. Work product created by any of the undersigned firms or agent of the undersigned firms that is related to this litigation, and that post-dates the filing of the initial complaint.

M. Timing of Privilege Logs: Privilege Logs shall be produced and/or supplemented on a rolling basis contemporaneously with each partial production of documents, and in no event later than thirty(30) calendar days after each such logged document is produced (or would have been produced, but for withholding or redaction).

N. Processing Specifications: The Producing Party will generate and preserve the MD5 or SHA-1 hash values of all ESI based on the Native Format file. Where Documents are not otherwise produced natively, all TIFF images shall display deletions, revisions, comments, speaker notes, and other rich data (including, but not limited to, hidden text, strikethrough text, the identity of the person making the deletion or revision and the date and time thereof, etc.) as displayed in the Document, regardless of the display setting for this information as last saved by the custodian, to the extent reasonably possible. Producing Parties shall make reasonable efforts to process all ESI with a single time zone and a date and time setting that is consistent across all of the Parties' productions. The Parties agree to meet and confer in good faith with regard to any issues arising from this aspect of production.

O. Production Media: The Producing Party shall produce Document images, Native Format files, load files, and Metadata via secure file transfer protocol (“FTP”) or physical media (“Production Media”). For physical media, each piece shall include a unique identifying label corresponding to the identity of the Producing Party (e.g., “DD”), the sequential production number (e.g., “PROD004”), and the Document Number ranges of the Documents in that production (e.g., “DD0000000123 - DD0000000456”). To the extent that the Production Media includes any confidential information protected under Protective Order, the label on such Production Media shall indicate that the Production Media includes information so designated as required under the terms of the Protective Order. Production Media shall include text referencing the case name and number. Further, any replacement Production Media shall cross-reference the original Production Media, clearly identify that it is a replacement, and cross-reference the BEGDOC and ENDDOC range that is being replaced. All Production Media that is capable of write protection should be write-protected before production. All Production Media may be encrypted, with the Producing Party to provide a decryption key under separate cover at the time of production.

P. Third Party Documents and ESI

A Party that issues a subpoena upon any third-party (“Issuing Party”) shall include a copy of this Protocol and the Protective Order with the subpoena. The Issuing Party shall produce a copy to all other Parties of any ESI or hard copy (including any metadata) obtained under subpoena to a third party. If a third-party production is not Bates-stamped, the Issuing Party will endorse the non-Party production with unique Bates prefixes and numbering scheme prior to reproducing them to all other Parties.

Q. Non-English Language Documents and ESI

Case No. 1:21-cv-05807-LAK-VF

STIPULATED PROTOCOL FOR DISCOVERY OF ELECTRONICALLY STORED INFORMATION (“ESI PROTOCOL”)

1. Metadata: In addition to the required production Metadata and Load File formats specified elsewhere in this ESI Protocol, including in Exhibit A, the Metadata for any translations produced to another Party shall clearly and logically reflect the translated document's relationship to the original language document, in a consistent manner designed to facilitate automated identification of all translations produced that are associated with a given original language document. In addition, the primary language for each document produced shall be set forth in the accompanying Metadata to the extent reasonably practicable using automated processes. The Parties will meet and confer on the necessity of additional formatting processes to enhance the production of Non-English language Documents and ESI, including but not limited to: (a) additional or different ESI metadata and coding fields; and (b) alternative technical specifications for documents containing text or characters which may be rendered inaccurate or illegible through the imaging and/or OCR process(es).

2. OCR Text: Non-English language Hard Copy Documents shall be produced with corresponding searchable multi-page optical character recognition ("OCR") text. Settings such as "auto-skewing" and "auto-rotation" should be turned on during the OCR process. All OCR shall be performed in a manner that takes into account the primary language contained in a document, so as to render the most useful resulting OCR text.

3. Search Terms: For any search terms to be utilized in this action, the Parties agree to meet and confer in good faith in order to arrive at mutually agreeable translations of such search terms into any relevant languages, dialects, or character sets. One or more

of the original and any translated forms of any search term may be applied to search a Party's ESI, as agreed by the Parties.

VIII. CROSS BORDER DISCOVERY

• ~~Defendants' Position~~

- ~~Defendants' Proposed Language:~~ The parties (including PRC counsel) will work cooperatively to address any complications of PRC discovery as efficiently as possible.

~~Defendants' Explanation:~~ As Plaintiffs are aware, coordination between Plaintiffs' PRC counsel and Defendants' PRC counsel has been necessary and helpful in similar cases involving cross-border discovery with Chinese parties. Defendants see no reason why Plaintiffs would be unwilling to assist in the collection of the documents that they are seeking here so that they can obtain quicker access to these documents in a manner that would satisfy both United States and Chinese legal systems.

~~Plaintiffs' Position~~

- ~~Plaintiffs' Proposed Language:~~ [N/A. Plaintiffs believe this section is unnecessary.]
- ~~Plaintiffs' Explanation:~~ As detailed above in Section I.2, the parties will work cooperatively and in good faith during the discovery process. However, Plaintiffs have no role to play in "assist[ing] in the collection of [Defendants'] documents." Nor is it Plaintiffs' responsibility to engage PRC counsel to advise Defendants on compliance with PRC law. To the extent Defendants are faced with "complications" arising from conflicts between PRC law and US law

Case No. 1:21-cv-05807-LAK-VF

(including this Court's orders), it is solely Defendants' responsibility to decide whether they comply with US law or PRC law. The "similar case" to which Defendants refer is a cautionary tale. In *In re NIO Inc. Securities Litigation*, Case No. 1:19-cv-01424-NGG-JRC (E.D.N.Y.), despite plaintiffs' counsel's attempt to assist defendants by engaging PRC counsel, the PRC-based defendant has stalled discovery for nearly three years. In *NIO*, the court denied the defendants' motion to dismiss in August 2021 and ordered the completion of fact discovery by December 2022. After repeated excuses, delays, and extensions, defendant NIO has produced just 68 documents to date—while claiming that thousands of documents are still in the process of being approved by PRC authorities—and fact discovery is still ongoing. *See, e.g., NIO* Dkt No. 135 (Letter from Robert Fumerton, dated May 16, 2024) ("While additional evidence of construction is in NIO's possession, it cannot be produced or relied on here until NIO receives authorization from the relevant Chinese authorities. NIO has been working diligently to obtain such authorization, in consultation with Plaintiffs' U.S. and PRC counsel."). Plaintiffs have no interest in allowing Defendants in this action to attempt to shift the burden of any such similar delays onto Plaintiffs.

IX. MODIFICATION

This protocol may be modified by written agreement of the Parties or by the Court for good cause shown.

IT IS SO STIPULATED, THROUGH COUNSEL OF RECORD.

Respectfully submitted,

Dated: May 31, 2024

/s/ Corey Worcester

Corey Worcester
Renita Sharma
Quinn Emanuel Urquhart & Sullivan, LLP
51 Madison Avenue, 22nd Floor
New York, New York 10038
Tel: (212) 849-7000
Fax: (212) 849-7100
coreyworcester@quinnemanuel.com
renitasharma@quinnemanuel.com

/s/ Phillip Kim

Phillip Kim
Laurence Rosen
The Rosen Law Firm, P.A.
275 Madison Avenue, 40th Floor
New York, New York 10016
Tel: (212) 686-1060
Fax: (212) 202-3827
pkim@rosenlegal.com
lrosen@rosenlegal.com

*Counsel for Defendants DiDi Global Inc.,
Will Wei Cheng, Jean Qing Liu, Stephen
Jingshi Zhu, Alan Yue Zhuo, and Adrian
Perica*

Lead Counsel for Lead Plaintiff and the Class

/s/ Jonathan Rosenberg

Jonathan Rosenberg
Abby F. Rudzin
O'MELVENY & MYERS LLP
7 Times Square
New York, NY 10036
Tel: (212) 326-2000
jrosenberg@omm.com
arudzin@omm.com

Scott Musoff
Robert Fumerton
Michael Griffin
Skadden, Arps, Slate, Meagher & Flom LLP
One Manhattan West
New York, New York 10001
Tel: (212) 735-3902
Fax: (212) 777-3902
smusoff@skadden.com
robert.fumerton@skadden.com
michael.griffin@skadden.com

*Counsel for Defendants Goldman Sachs
(Asia) L.L.C.; Morgan Stanley & Co. LLC;
J.P. Morgan Securities LLC; BofA Securities,
Inc.; Barclays Capital Inc.; Citigroup Global
Markets Inc.; HSBC Securities (USA) Inc.;
UBS Securities LLC; Mizuho Securities USA
LLC; and China Renaissance Securities
(US) Inc.*

Counsel for Defendant DiDi Global Inc.

/s/ Sheryl Shapiro Bassin

Sheryl Shapiro Bassin
WILSON SONSINI GOODRICH &
ROSATI, P.C.
1301 Avenue of the Americas, 40th Floor

/s/ Jeffrey T. Scott

Jeffrey T. Scott
Andrew M. Kaufman
Sullivan & Cromwell LLP
125 Broad Street
New York, New York 10004-2498
Tel: (212) 558-3082
Fax: (212) 291-9138
scottj@sullcrom.com
kaufmana@sullcrom.com

Counsel for *Defendant Kentaro Matsui*

/s/ Matthew S. Kahn

Matthew S. Kahn
One Embarcadero Center
Suite 2600
San Francisco, California 94111
Telephone: (415) 393-8200
Facsimile: (415) 393-8306
mkahn@gibsondunn.com

Michael D. Celio (admitted *pro hac vice*)
Kevin J. White
310 University Avenue
Palo Alto, California 94301
Telephone: (650) 849-5300
Facsimile: (650) 949-5333
mcelio@gibsondunn.com
kwhite@gibsondunn.com

Counsel for *Defendant Martin Chi Ping Lau*

New York, New York 10019

Tel: (212) 999-5800
Fax: (212) 999-5899
SBassin@wsgr.com

Ignacio E. Salceda (admitted *pro hac vice*)
WILSON SONSINI GOODRICH &
ROSATI, P.C.
650 Page Mill Road
Palo Alto, California 94304
Tel: (650) 493-9300
Fax: (650) 493-6811
ISalceda@wsgr.com

Counsel for *Defendant Zhiyi Chen*

/s/ Stephen P. Blake

Stephen P. Blake
Bo Bryan Jin
SIMPSON THACHER & BARTLETT LLP
2475 Hanover Street
Palo Alto, CA 94304
Telephone: (650) 251-5000
Facsimile: (650) 251-5002
sblake@stblaw.com
bryan.jin@stblaw.com

Counsel for *Defendant Daniel Yong Zhang*

SO ORDERED


LEWIS A. KAPLAN, USBJ

6/11/24

SO ORDERED.

DATED: _____

Hon. Lewis A. Kaplan
U.S. District Judge

CERTIFICATE OF SERVICE

IT IS HEREBY CERTIFIED that on this 31st day of May 2024, this document filed through the ECF system will be sent electronically to the registered participants on the Notice of Electronic Filing and paper copies will be sent to any non-registered participants.

/s/ Laurence Rosen